Population Ethics under Risk*

Gustaf Arrhenius†   H. Orri Stefánsson‡

July 8, 2020

Abstract

Population axiology concerns how to rank populations by the relation "is socially preferred to". So far, this field has largely ignored issues about uncertainty and the axioms that have been discussed mostly pertain to the ranking of risk-free outcomes. Most public policy choices, however, are decisions under uncertainty, including policy choices that affect the size of a population (such as climate policy choices). Here, we shall address the question of how to rank population prospects—that is, alternatives that contain uncertainty as to which population they will bring about—by the relation "is socially preferred to". We start by illustrating how well-known population axiologies can be extended to population prospect axiologies. And we show that new problems arise when extending population axiologies to prospects. In particular, traditional population axiologies lead to prospect-versions of the problems that they are praised for avoiding in the risk-free settings. Finally, we identify an axiom that, we contend, should be satisfied by any population prospect axiology, and show how given this axiom, the impossibility theorems in population axiology can be extended to impossibility theorems for *non-trivial* population prospects, that is, prospects that confer probabilities *strictly between* zero and one on different populations.

†**Email: gustaf.arrhenius@iffs.se** Institute for Futures Studies, Stockholm; Department of Philosophy, Stockholm Univeresity.

‡**Email: orri.stefansson@philosophy.su.se.** Department of Philosophy, Stockholm University; Swedish Collegium for Advanced Study, Uppsala; Institute for Futures Studies, Stockholm.

1   Introduction

Population axiology concerns how to evaluate populations in terms of what philosophers call their "moral goodness", that is, how to rank populations by the relation "is socially preferred to", to use terminology more familiar in social choice theory and welfare economics. So far, discussions of population axiology have largely ignored the question of how to rank alternatives that contain uncertainty as to which population they will bring about.[1] The adequacy conditions (i.e., "axioms") that have been discussed mostly pertain to the ranking of risk-free outcomes and are silent in regard to how to rank (non-trivial) risky prospects. In other words, questions about what we call *population prospect axiology*—that is, questions about how to rank what we call *population prospects*—have not yet been much addressed.

This exclusion of uncertainty is unfortunate since most public policies that could affect the number and identities of future people contain a great deal of uncertainty as to which population they will bring about.[2] For instance, as the Intergovernmental Panel on Climate Change acknowledged in their Fifth Assessment Report,[3] we should expect that the climate policies we choose will affect the size of the world population. But obviously, climate policies contain a great deal of uncertainty, for instance, uncertainty about what type of population they will bring about.[4]

Apart from the need to consider population prospects for policy purposes, there is a theoretical motivation for introducing uncertainty into population ethics. One might hope that removing the idealizing assumption of certainty could help us avoid the impossibility results that have been troubling population ethicists for decades.[5] After all, removing information can result in a situation where an adequacy condition that is used in an impossibility theorem no longer applies. So, removing information that is necessary for a condition to apply could be seen as one possible way in which one can avoid

---

[1]For important exception, see Blackorby et al. (2005), Roberts (2007), Asheim and Zuber (2016), Nebel (2019), Budolfson and Spears (2018), Spears and Budolfson (2019). Voorhoeve and Fleurbaey (2016) and Nebel (2017) discuss prospects involving populations with different people but the same number of people.

[2]Contrary to the convention in economics, we use "uncertainty" and "risk" interchangeably throughout this paper. So, a situation of risk, and a situation of uncertainty, is one where one knows what outcome one's choices *might* result in, but does not know which outcome one's choice *will* result in.

[3]Available at `https://www.ipcc.ch/assessment-report/ar5/`.

[4]For further examples, see Broome (1992, 2004, 2010, 2015). See also Arrhenius (forthcoming, ch. 1) for a discussion.

[5]See e.g., Parfit (1984); Ng (1989), Arrhenius (2000a), Arrhenius (2000b); Blackorby et al. (2005), Arrhenius et al. (2017).

an impossibility result involving that condition.[6] In fact, that is what happens when one introduces risk and uncertainty into population ethics—that is, when one removes information about which state of the world is actual—since the traditional adequacy conditions for population axiology say nothing about how to rank what we call *non-trivial* prospects, that is, prospects that confer probabilities *strictly between* zero and one on different populations (as we further explain below).

An axiology for non-trivial population prospect that avoids impossibility results would moreover be a great victory from a practical point of view. Policies that could affect the size of a population typically contain considerable uncertainty, as previously mentioned. In fact, it is hard to think of any policy (population-affecting or not) that is certain to deliver a particular outcome. Therefore, one might conjecture that from a practical point of view, the impossibility results for risk-free population axiology are not so important. What is important from a practical point of view, is the possibility of an adequate axiology for non-trivial population prospects. For if we have an adequate population prospect axiology, then we will be able to make consistent population policy choices (since these are always choices between non-trivial prospects), notwithstanding the traditional impossibility results for risk-free populations.

Another reason for exploring what happens when we introduce probabilities in population ethics, is that we know that there are some multi-person decision problems where a solution can be found only if we introduce probabilities. In particular, there are some well-known problems in game theory where a *Nash equilibrium* exists only if the "players" are allowed to randomise between their available "pure" (i.e., non-probabilistic) strategies (von Neumann and Morgenstern, 1944). In a similar vein, one might hope that a solution can be found to the problems of population ethics by introducing probabilities.

Unfortunately, it turns out that new problems appear when one tries to turn a population axiology into a population prospect axiology. Perhaps most surprisingly, we shall see that natural formulations of traditional (risk-free) population axiologies lead to prospect-versions of the problems that they are praised for avoiding in the risk-free setting.

Moreover, we shall identify an intuitive adequacy condition that, we contend, should be satisfied by any social preference ordering of population prospects, but which opens the door to impossibility results for population prospect axiology that are analogous to the traditional impossibility results for (risk-free) population axiology. Importantly, these

---

[6]Cf. Arrhenius's (forthcoming) discussion (in section 6.4) of how restriction on the measurement assumptions regarding welfare can help avoiding certain impossibility results in non-risky population axiology.

results cannot be derived just from the (risk-free) adequacy conditions involved in the traditional impossibility theorems, and they hold for *non-trivial* prospects (as well as for trivial ones), that is, for prospects that confer probabilities strictly between zero and one on different populations. So, contrary to the aforementioned hope, removing the idealizing assumption of certainty will unfortunately not help us avoid the notorious impossibility results in population ethics.

2   Framework and Background Assumptions

A natural starting point in the search for a population prospect axiology is to consider some extensions of standard population axiologies to situations of uncertainty. But first, we need to introduce a few terms and distinctions used in traditional, risk-free, population axiology, which we will use when formulating the aforementioned extensions.

In traditional population axiology, one distinguishes between the welfare of a person and the contributive value of a person's welfare (or well-being; we shall use these two terms interchangeably). By the contributive value of a person's welfare we shall mean the value that a person's welfare contributes to the value of a population of which it is a member. More exactly, the contributive value of a person i's welfare relative to a population $\mathbf{x}$, of which i is a member, is the difference in value between $\mathbf{x}$ and the population consisting of all the $\mathbf{x}$-people except i. This distinction is useful for population prospect axiology too, for we can now reformulate one of our main questions as "What is the contributive value of a possible person's welfare to a prospect?"

Now to our formal framework, which follows closely that of Asheim and Zuber (2014), although suitably extended to prospects based on the framework of Savage (1954). Let $\mathbb{N}$ denote the set of natural numbers and $\mathbb{R}$ the set of real numbers. $\mathbb{R}^- \subset \mathbb{R}$ and $\mathbb{R}^+ \subset \mathbb{R}$ are respectively the set of negative and positive real numbers. Let $\mathbf{X} = \bigcup_{n \in \mathbb{N}} \mathbb{R}^n$ denote the set of possible finite distributions of lifetime well-being. More formally, $\mathbf{X} = \bigcup_{n \in \mathbb{N}} \mathbb{R}^n$ is a set of of vectors of real numbers, where each number ("utility") represents the lifetime well-being of some person. A generic such vector for a population of m people is denoted $\mathbf{x} = (x_1, ..., x_m)$, where x_i denotes the lifetime well-being of individual i. Informally, we shall talk of a vector like this as a *population*. $\mathbf{X}^- \subset \mathbf{X}$ and $\mathbf{X}^+ \subset \mathbf{X}$ are the subsets of such vectors where *every* number in the vector is negative respectively positive. The size of the population given by $\mathbf{x}$ is denoted by $n(\mathbf{x})$ (and will, as mentioned, always be finite). For any vector $\mathbf{x}$, we write the average value of its elements as $\bar{u}$. So, $\bar{u}$ should be interpreted

as the average lifetime welfare of people in population $\mathbf{x}$.

Let $(z)_n \in \mathbb{R}^n$ denote the perfectly-equal population where all n individuals have lifetime well-being z. We let $(\mathbf{x}, \mathbf{y})$ denote the population that results when population $\mathbf{x} \in \mathbf{X}$ is conjoined with population $\mathbf{y} \in \mathbf{X}$. And let $(\mathbf{x}, (z)_n)$ denote population $\mathbf{x} \in \mathbf{X}$ with n added individuals that all have lifetime well-being z.

The well-being measure is normalised around the "neutral" life, such that if i's life in $\mathbf{x}$ is neutral, then $x_i = 0$, while if i's life in $\mathbf{x}$ is better for i than a neutral life, then i has positive welfare in $\mathbf{x}$, i.e., $x_i > 0$, but if i's life in $\mathbf{x}$ is worse for i than a neutral life neutral, then i has negative welfare in $\mathbf{x}$, i.e., $x_i < 0$. We shall remain agnostic about what constitutes well-being and what determines whether a life is neutral (but for different suggestions for the latter, see e.g. Arrhenius, forthcoming, ch. 2 and 9, Arrhenius, 2000a, Broome, 2004, and Parfit, 1984, pp. 357-358, and appendix G). Importantly, we do not need an analysis of a neutral welfare in the present context but rather just a criterion, and the criterion can vary with different theories of welfare.

To simplify the formal presentation of views and conditions, we shall assume the standard *anonymity* assumption, which holds that the "social welfare relations" that we study are invariant under permutations of the vectors in $\mathbf{X}$. So, for instance, let $\mathbf{x}'$ be the vector that results when the lifetime well-being of i and j in $\mathbf{x}$ are switched. Then the social welfare relations that we shall consider are all indifferent between $\mathbf{x}$ and $\mathbf{x}'$, that is, they deem these two populations to be equally good. Intuitively, this means that it does not matter *who* receives what welfare; all that matters is how welfare quantities are distributed. This assumption rules out some views in the literature that are called "person affecting" (cf. Glover, 1977) as well as desert-based views (Feldman, 1997; Arrhenius, 2003a). But we should emphasise that anonymity is just a simplifying assumption. Just like the original results in the risk-free setting, the results of this paper can be established without an anonymity assumption, but the argument would then involve more complicated formulations.

We let $\precsim$ on $\mathbf{X}$ denote a "social welfare relation" on $\mathbf{X}$, such that for any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, $\mathbf{x} \precsim \mathbf{y}$ means that $\mathbf{y}$ is at least as socially preferable as $\mathbf{x}$, or, as we shall often put it, $\mathbf{y}$ is at least as *good* as $\mathbf{x}$. (So, $\precsim$ is what some philosophers would call a "moral goodness" relation, but will shall from now on refer to it as a social welfare relation.) To simplify the discussion we shall typically assume that the social welfare relation is transitive, reflexive, and complete, which means that the relation generates a social welfare *order*. It is worth noting, however, that for the impossibility results that we present in section 5 a *quasi*-order suffices, that is,

completeness need not be assumed. In other words, these results are consistent with some populations being incommensurable in value.[7] The strict relation, $\prec$, and indifference, $\sim$, are respectively the asymmetric and symmetric counterparts of $\precsim$. When we speak of a "social planner", we mean the (possibly group) agent whose preference is denoted by $\precsim$.

Although we shall typically assume that the relevant probabilities are "objective", or given to the social planner, as in the von Neumann and Morgenstern (1944) framework, we use a more general Savage (1954)-inspired framework—which does not assume the probabilities to be given to the decision-maker—to turn the traditional population axiologies into population prospect axiologies. To that end, let $\mathbf{S}$ be a set of (non-overlapping) *states of the world*, that is, resolutions of uncertainty, with generic elements s_i, s_j. For simplicity, we assume $\mathbf{S}$ to be finite. Let $\mathbf{P}$ be a set of functions from $\mathbf{S}$ into $\mathbf{X}$. Generic elements of $\mathbf{P}$ will be denoted P, Q, or P_1, P_2, etc. We interpret the elements in $\mathbf{P}$ as *population prospects*. $P(s_i) \in \mathbf{X}$ is the well-being vector, that is, the population, that results from prospect P if s_i is the actual state of the world. Finally, we assume that the social planner has a probability measure over $\mathbf{S}$, that is, there is a function p from $\mathbf{S}$ into [0,1] such that $\sum_{i=1} p(s_i) = 1$, which quantifies the social planner's uncertainty (which may but need not correspond to some objective probability). To keep the notation as simple as possible, we also use $\precsim$ for the social planner's preference relation over $\mathbf{P}$.

It will be useful to work with *events*, that is, unions of states, when stating some conditions. To that end, let $\mathbf{E}$ be a Boolean algebra of events generated from $\mathbf{S}$, that is, the set of all possible sets of states in $\mathbf{S}$, closed under the Boolean operators. Generic elements of $\mathbf{E}$ will be denoted E, $\neg E$, F, etc. Since we assume the social planner to be probabilistically coherent, her probability function extends straightforwardly from $\mathbf{S}$ to $\mathbf{E}$. Moreover, probabilistic coherence ensures that if $E \subset \mathbf{E}$ then $p(E) < 1$ and $p(\neg E) > 0$, which will be convenient when stating some of the below conditions. Finally, to further simplify notation, when we for instance write $P(E) = \mathbf{x}$, this means that *every* state in E results in $\mathbf{x}$ given P, that is, for any $s_i \in E$, $P(s_i) = \mathbf{x}$.

Later we shall focus much of our attention on what we call *non-trivial* prospects. Informally, a prospect is non-trivial if it confers a positive probability on at least two distinct populations; otherwise it is said to be trivial. More formally:

Definition (Non-trivial prospect). *A prospect, $P \in \mathbf{P}$ is called* non-trivial *just in case there are $P(s_i), P(s_j) \in \mathbf{X}$ where $P(s_i) \neq P(s_j)$ and $p(s_i) > 0$, $p(s_j) > 0$.*

---

[7]For recent discussions of limited, or "imprecise", comparability in the context of population ethics, see Arrhenius (2016) and Parfit (2016). See also Blackorby et al. (2005) on incomparability between populations.

We let $\mathbf{P}_{nt} \subset \mathbf{P}$ denote the set of non-trivial prospects.

3   Population Prospect Axiologies

Let us now consider how the traditional population axiologies could be extended to situations of uncertainty, using the the framework and terminology from the last section.[8] We start with Total Utilitarianism, according to which the contributive value of a person's welfare equals her welfare,[9] and the value of a population is calculated by summing the welfare of all people in the population.

Total Utilitarianism (TU). *For any $x, y \in X$:*

$$x \precsim y \Leftrightarrow TU(x) = \sum_i x_i \leq TU(y) = \sum_i y_i$$

A natural way to turn Total Utilitarianism into a population prospect axiology is to sum up weighted Total Utilitarian values of all the possible populations that a prospect might result in, where the weight on the value of each population is determined by the probability that the prospect will result in that population.[10] Let us call the resulting axiology *Prospect Total Utilitarianism.*

---

[8]To focus the discussion—and due to limited space—we shall mostly ignore rank-dependent views, and other views that have been called "variable-value views", early version of which were discussed by Hurka (1983), Ng (1989), Sider (1991). More recent discussions of such views include Zuber and Asheim (2012), Asheim and Zuber (2014, 2016), and Pivato (2020).

[9]Strictly speaking, the contributive value of a person's welfare can be a linear transformation of her welfare, according to Total Utilitarianism; but it can neither be a concave nor a convex transformation of her welfare.

[10]The population prospects axiologies we formulate treat risk and uncertainty as standard expected utility theory does. There are reasons for being sceptical of that treatment; for instance, since it does not allow for risk aversion with respect to utility. However, to keep things simple, we shall assume the standard expected utility treatment of risk. It is easy to see that the results we shall discuss would still hold if we assumed some of the minimally normative alternative treatments of risk that have recently been suggested, that is, treatments that satisfy transitivity and stochastic dominance, such as the theories suggested by Quiggin (1982), Tversky and Kahneman (1992), Buchak (2013) and Stefánsson and Bradley (2015, 2019). In contrast, our results do not hold if one uses, say, the original Prospect Theory (Kahneman and Tversky 1979), which cannot simultanously satisfy stochastic dominance and transitivity, or Regret Theory (Bell, 1982; Loomes and Sugden, 1982), which only satisfies transitivity w.r.t. a triple of alternatives if at least one alternative in the triple strictly state-wise dominates another alternative in the triple (Diecidue and Somasundaram, 2017). While the latter theories may be accurate descriptions of how people actually rank prospects, few if any scholars would suggest them as theories about how one ideally should order prospects. Therefore, we think that we can justifiably set them aside for the present purposes, where the aim is to discuss how one *should* rank population prospects.

Prospect Total Utilitarianism (PTU). *For any $P, Q \in \mathbf{P}$:*

$$P \precsim Q \Leftrightarrow \sum_i p(s_i)TU(P(s_i)) \leq \sum_i p(s_i)TU(Q(s_i))$$

One of the great challenges that examinations of different size populations have raised for Total Utilitarianism, is how to respond to the fact that the latter gives rise to what Parfit (1984) called the *Repugnant Conclusion*. Informally, this conclusion is that for any perfectly equal population (no matter how well off are its members), there is a better population consisting of people with very low positive welfare. More formally:[11]

Repugnant Conclusion. *For all $y, z \in \mathbb{R}$, where $y > z > 0$, and for any $k \in \mathbb{N}$, there is a $n \in \mathbb{N}$, where $n > k$, and such that $(y)_k \prec (z)_n$.*

It should be evident that the prospect version of Total Utilitarianism also leads to the Repugnant Conclusion not only in non-risky situations but also in risky ones. For instance, for any prospect P, all of whose possible outcomes are populations where people enjoy very high welfare, Prospect Total Utilitarianism will recommend choosing instead some prospect Q, all of whose possible outcomes are larger populations where people have very low positive welfare. The result, if we follow the recommendation of Prospect Total Utilitarianism, is that we end up with a large population where people have very low positive welfare when we had the option of some smaller population where people enjoy very high welfare—an instance of the Repugnant Conclusion. Likewise, for other potentially counterintuitive implications of Total Utilitarianism in non-risky settings, such as the Very Repugnant Conclusion.[12]

Moreover, Prospect Total Utilitarianism violates probabilistic analogues of the traditional (risky-free) Quality condition which essentially rules out the Repugnant Conclusion (and which we state in section 5). Let $\mathbb{H} \subset \mathbb{R}^+$ be welfare levels that are very high—such that a life at a level in $\mathbb{H}$ would be deemed excellent—while $\mathbb{L} \subset \mathbb{R}^+$ are welfare levels that

---

[11]It is worth noting that this formulation is more general than Parfit's (1984): "For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better even though its members have lives that are barely worth living" (388). Since the reason we discuss the Repugnant Conclusion in this article is only to illustrate what some have taken to be an undesirable implication of Total Utilitarianism, no harm is done by focusing on the more general version, which Total Utilitarianism of course implies.

[12]Informally, the Very Repugnant Conclusion states that for any perfectly equal high-welfare population, and for any number of people with very negative welfare, there is a population consisting of the people with negative welfare and people with very low positive welfare which is better than the high welfare population (Arrhenius, 2003b).

are very low but positive—such that a life at a level in $\mathbb{L}$ would be deemed barely worth living. Consider the following probabilistic analogue of the traditional Quality condition, which can be derived from the non-risky condition with help of a dominance condition that we discuss in section 4, but which is essentially the traditional State-Wise Dominance principle from decision theory. (Recall that $P(E) = \mathbf{x}$ means for any $s_i \in E$, $P(s_i) = \mathbf{x}$; hence, we can apply State-Wise Dominance when, say, $P(E) = \mathbf{x}$ and $P(\neg E) = (z)_n$.) Informally, what we shall call *Probabilistic Quality* says that there is some very high welfare level z, such that for any very low but positive welfare level y, if prospects P and Q result in the same population in some state(s), but in the state(s) where they result in different populations, P results in a population where every person is at the very high level z whereas Q results in a population where every person is at the very low level y, then P is at least as good as Q. More formally:

Probabilistic Quality. *There is a $z \in \mathbb{H}$ and a $n \in \mathbb{N}$, such that for any $\mathbf{x} \in \mathbf{X}$ and any $E \subseteq \mathbf{E}$ and any $y \in \mathbb{L}$ and any $m \in \mathbb{N}$, if $Q(E) = \mathbf{x}$ and $Q(\neg E) = (z)_n$, while $P(E) = \mathbf{x}$ and $P(\neg E) = (y)_m$, then $P \precsim Q$.*

Prospect Total Utilitarianism of course violates Probabilistic Quality. The expected utility framework we have assumed when extending Total Utilitarianism to Prospect Total Utilitarianism implies that the probability and utility of $\mathbf{x}$ can be ignored when comparing P with Q (this follows from the aforementioned dominance condition). Hence, since m can be very large, compared with n, Prospect Total Utilitarianism violates Probabilistic Quantity, for the same reason that Total Utilitarianism leads to the Repugnant Conclusion.

In comparison with other axiologies, Total Utilitarianism handles some (non-risky) cases involving negative welfare better than most alternative axiologies, such as Average and Critical-Level Utilitarianism (more on which below). For example, Total Utilitarianism clearly avoids the Very Sadistic Conclusion which is implied by Critical-Level Utilitarianism.[13] Informally, the Very Sadistic Conclusion is that for any population with negative welfare, there is a population with positive welfare which is worse. More formally:

Very Sadistic Conclusion. *For any $\mathbf{x} \in \mathbf{X}^-$ there is a $\mathbf{y} \in \mathbf{X}^+$ such that $\mathbf{y} \prec \mathbf{x}$.*

While Total Utilitarianism avoids the Very Sadistic Conclusion, Prospect Total Utilitarianism leads to what we call the *Risky Very Sadistic Conclusion*.[14]

---

[13]For a discussion, see Arrhenius (2000a,b, forthcoming). Average Utilitarianism has similar problems.

[14]We note that Budolfson and Spears (2018) discuss a more general result, which they call the Probabilistic Very Repugnant Conclusion. A proper comparison of the results we discuss here to the results discussed by Budolfson and Spears will have to await another occasion.

Risky Very Sadistic Conclusion. *For any $E \subset \mathbf{E}$, for any $\mathbf{x} \in X^-$, for any $\mathbf{y} \in X^+$, and for any $\gamma \in \mathbb{R}^+$, there is a $z \in \mathbb{R}$ such that $z < \gamma$ and a $m \in \mathbb{N}$, and such that if $P(E) = \mathbf{x}$ and $P(\neg E) = (z)_m$, while the* trivial *prospect Q for sure results in $\mathbf{y}$, then $Q \prec P$.*

Here is why Prospect Total Utilitarianism implies the Risky Very Sadistic Conclusion. For any probability $p(E) < 1$, however high, if m is sufficiently great, then the expected total welfare of prospect P, which has probability p of resulting in $\mathbf{x}$ and probability $1 - p$ of resulting in $(z)_m$, will be higher than the expected total welfare of prospect Q, which results in $\mathbf{y}$ for sure. Thus, Prospect Total Utilitarianism deems P better than Q. This holds irrespective of how much the $\mathbf{x}$-people suffer and of how many they are. In other words, while Total Utilitarianism avoids the Very Sadistic Conclusion, Prospect Total Utilitarianism recommends prospects that will almost certainly result in a population with negative welfare even though a population with positive welfare was available (hence the name of the conclusion).[15]

Now, there is of course a sense in which the above conclusion is just an instance of a general implication of Expected Utility Theory (EUT)—and, in particular, its Continuity axiom—that has nothing specifically to do with population ethics.[16] However, we think that while one might see the Risky Very Sadistic Conclusion as a special case of a more general implication of EUT, that does not change the fact that the special case in question is illuminating for debates in population ethics. In particular, it teaches us something important about Total Utilitarianism, namely, that while the theory is praised for avoiding the Very Sadistic Conclusion, the most natural prospect-extension of the theory comes arbitrarily close to the very same conclusion, in the sense of recommending, for any population with negative welfare, a choice that has an arbitrarily high probability of resulting in that population, even though a population with positive welfare was available. So, the theory comes arbitrarily close to recommending the Very Sadistic Conclusion.[17]

---

[15]Some may dislike the name of this conclusion for the following reason: Total utilitarians think that the population with negative welfare is the worst of the possible populations involved in the above prospect; hence, the Risky Very Sadistic Conclusion differs in an important respect from the ordinary Very Sadistic Conclusion, where critical-level utilitarians think that a population with negative welfare is *better* than one with positive welfare. Nevertheless, we think that the label is apt for the reason explained above: The Very Risky Sadistic Conclusion will almost certainly result in an outcome analogous to the outcome of the Very Sadistic Conclusion.

[16]We thank Johan Gustafsson for pressing us on this point.

[17]We also note that this implication of Prospect Total Utilitarianism, as well as the implication of Prospect Average Utilitarianism that we discuss below, supports the suggestions by Spears and Budolfson (2019) and Budolfson and Spears (2018) that both "repugnance" and "sadism", given their general understanding of these terms, is harder to avoid than what many scholars seem to think (as also pointed out by Arrhenius 2000a, forthcoming).

Faced with the Repugnant Conclusion and other similar problems, a Utilitarian might be tempted to opt for a non-total version of the theory, such as Average Utilitarianism or Critical-Level Utilitarianism, which can respectively be extended to Prospect Average Utilitarianism and Prospect Critical Level Utilitarianism. It is easy to see that these theories avoid the Repugnant Conclusion. However, as is well known, these alternatives to Total Utilitarianism have, in the risk-free setting, their own counterintuitive conclusions to deal with; conclusions that are arguably at least as counterintuitive as the Repugnant Conclusion.[18] It should be evident that since Prospect Average Utilitarianism and Prospect Critical Level Utilitarianism are extensions of Average Utilitarianism and Critical Level Utilitarianism to risky situations, they inherit the problems of their ancestors for *trivial* prospects.[19] Moreover, it turns out that Prospect Average Utilitarianism and Prospect Critical Level Utilitarianism—and, in fact, any population prospect axiology that satisfies transitivity and a weak dominance condition—faces additional problems due to the introduction of uncertainty, even when the scope is limited to non-trivial prospects, as we shall soon see.

But first, to appreciate some of the problems that non-total Utilitarian theories are faced with in the risk-free setting, let's look more closely at Average Utilitarianism, according to which the value of a population is found by summing up the welfare of the people in the population and dividing this sum with the number of people in the population.[20] (After examining Average Utilitarianism and its prospect extension, we shall briefly discuss Critical Level Utilitarianism and its prospect extension.) Hence, the contributive value of a person's welfare depends on the average welfare of the population of which it is a member: The contributive value of a person's welfare is positive (negative, neutral) if the person has a higher (lower, same) welfare than the average welfare in the population.

Average Utilitarianism (AU). *For any $x, y \in X$:*

$$x \precsim y \Leftrightarrow AU(x) = \bar{u} \leq AU(y) = \bar{y}$$

Now, we can extend this theory to situations of risk—thus creating Prospect Average Utilitarianism—in an analogues way to how Total Utilitarianism was extended to Prospect

---

[18]See, e.g., Parfit (1984); Arrhenius (2000b, forthcoming).

[19]The same holds for the recent prospect axiology presented in Asheim and Zuber (2016) which is an extension of the population axiology in Asheim and Zuber (2014). The latter violates some compelling adequacy conditions, such as the Non-Extreme Priority Condition (Arrhenius, forthcoming).

[20]The problem with Average Utilitarianism that we focus on here also rises for Critical Level Utilitarianism. See Arrhenius (2000a,b, forthcoming).

Total Utilitarianism:

Prospect Average Utilitarianism (PAU). *For any $P, Q \in P$:*

$$P \precsim Q \Leftrightarrow \sum_i p(s_i) AU(P(s_i)) \leq \sum_i p(s_i) AU(Q(s_i))$$

Average Utilitarianism gives rise to the *Sadistic Conclusion*, which informally says that it is better to add to a population people with negative rather than positive welfare.[21] More formally:

Sadistic Conclusion. *There is a $x \in X$, a $y \in X^+$, and a $z \in X^-$ such that $(x, y) \prec (x, z)$.*

Even worse, Average Utilitarianism violates the seemingly unassailable *Weak Non-Sadism*, which is even logically weaker than what is required to avoid the Sadistic Conclusion. Informally, Weak Non-Sadism implies that there is a negative welfare level and a number of people at this level such that an addition of any number of people with positive welfare is at least as good as an addition of the negative welfare people. More formally:

Weak Non-Sadism. *There is a $z \in \mathbb{R}^-$ and a $n \in \mathbb{N}$, such that for any $x \in X$ and for any $y \in \mathbb{R}^+$ and for any $m \in \mathbb{N}$, $(x, (z)_n) \precsim (x, (y)_m)$.*

Prospect Average Utilitarianism of course violates Weak Non-Sadism too, that is, for trivial prospects. Moreover, it is straightforward to verify that Prospect Average Utilitarianism also violates, *for non-trivial prospects*, a probabilistic analogues of Weak Non-Sadism that is implied by Weak Non-Sadism together with the aforementioned dominance condition:

Probabilistic Weak Non-Sadism. *There is a $z \in \mathbb{R}^-$ and a $n \in \mathbb{N}$, such that for any $x, x' \in X$ and for any $y \in \mathbb{R}^+$ and for any $m \in \mathbb{N}$, and for any $P, Q \in P$, and for any $E \subseteq E$, if $P(E) = (x, (z)_n)$ and $P(\neg E) = (x, x')$, while $Q(E) = (x, (y)_m)$ and $Q(\neg E) = (x, x')$, then $P \precsim Q$.*

Perhaps more interestingly, while Average Utilitarianism (and Critical Level Utilitarianism) of course avoids the risk-free Repugnant Conclusion, Prospect Average Utilitarianism (and Prospect Critical Level Utilitarianism as well as of course Prospect Total Utilitarianism) leads to what we might call the *Risky Repugnant Conclusion*:

---

[21]See, e.g., Arrhenius (2000b).

Risky Repugnant Conclusion. *For any $E \subset \mathbf{E}$, for any $\mathbf{y}, \mathbf{z} \in \mathbf{X}$, there is a $\mathbf{x} \in \mathbf{X}$, such that if $P(\neg E) = \mathbf{x}$ and $P(E) = \mathbf{z}$ while trivial prospect Q results in $\mathbf{y}$ for sure, then $Q \prec P$.*

The above observation means that while Average Utilitarianism (and Critical Level Utilitarianism) avoids the Repugnant Conclusion, *Prospect* Average Utilitarianism (and Prospect Critical Level Utilitarianism) recommends prospects that will *almost certainly* result in a Repugnant Conclusion, that is, in the special case where **z** is a huge population consisting of people who have very low positive welfare while **y** is a smaller population where everyone is very well-off. So, again we see that a prospect axiology leads to the probabilistic analogue of the conclusion that the corresponding risk-free axiology is celebrated for avoiding.

For the sake of completeness, and to convince the reader that Prospect Critical Level Utilitarianism really does lead to the Risky Repugnant Conclusion, it is worth formally stating the theory. The general intuition behind Critical Level Utilitarianism is that only people with welfare that is above some "critical level", which is typically assumed to be higher than the neutral level, can increase the value of a population. More precisely, the risk-free version of the theory states that for some critical level $c \geq 0$:[22]

Critical-Level Utilitarianism (CLU)**.** *For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$:*

$$\mathbf{x} \precsim \mathbf{y} \Leftrightarrow CLU(\mathbf{x}) = \sum_i [u_i - c] \leq CLU(\mathbf{y}) = \sum_i [u_i - c]$$

Blackorby et al. (2005) have extended Critical Level Utilitarianism[23] to a prospect theory by essentially applying expected utility theory to population prospects where each population is evaluated by the above formula. That is, they extend Critical Level Utilitarianism in the same way as we have extended Average and Total Utilitarianism (albeit using a somewhat different notation and formal framework than we do). The resulting theory states that:

Prospect Critical-Level Utilitarianism. *For any $P, Q \in \mathbf{P}$:*

$$P \precsim Q \Leftrightarrow \sum_i p(s_i) CLU(P(s_i)) \leq \sum_i p(s_i) CLU(Q(s_i))$$

---

[22]It is worth noting that c need not be a constant, but can be a function of the size of the evaluated population. A generalisation of Critical Level Utilitarianism was introduced by Blackorby and Donaldson (1984) and subsequently explored in depth by Blackorby et al. (2005) and Bossert (2017). The generalisation, which we ignore, adds a function on $u_i - c$ that is non-decreasing and non-convex (so, weakly or strictly concave or linear).

[23]That is, the generalised version mentioned in footnote 22.

It should be evident that while Critical Level Utilitarianism was partly designed to avoid the Repugnant Conclusion, Prospect Critical Level Utilitarianism leads to the Risky Repugnant Conclusion. For as long as the people in the **x** population are sufficiently high above the critical level, then Prospect Critical Level Utilitarianism will instruct us to gamble on **x** rather than getting **y** for sure, irrespective of how much more likely the gamble is to result in **z** than in **x** and irrespective of how much worse off people are in **z** compared to in **y**.

Now, average and critical level utilitarians could avoid the Risky Repugnant Conclusion by postulating that there is a limit to how high a person's welfare can be. But what we take this and the previous result to illustrate—and the point we want to emphasise—is that new problems appear when we try to turn a population axiology into a population prospect axiology. For, as we have seen, natural formulations of each of the three standard versions of Utilitarianism lead to prospect-versions of the conclusions that they avoid in the risk-free settings.[24]

For the remainder of this paper, the question we will focus on is whether the impossibility results that have been proven for risk-free population axiology—that is, results showing that no population axiology can satisfy different sets of intuitively compelling adequacy conditions—also hold for non-trivial prospects. We know from the study of risk-free population axiology that there is a great advantage to focusing on exploring the logical consistency of different adequacy conditions (i.e., axioms), rather than formulating new theories in response to the deficiencies that have been found in the old theories. After all, we know, from some notorious impossibility results (reviewed in Arrhenius, forthcoming) that any (risk-free) population axiology will have some deficiency.

In the spirit of the axiomatic approach, we shall, in the next section, introduce a weak and very plausible condition that ensures that the ranking of population prospects coheres with the ranking of populations. As we show in section 5, this condition suffices to generate impossibility results for population prospect axiology that are similar to the traditional impossibility results for (risk-free) population axiology. Moreover, as we shall explain in section 5, something like the condition introduced in the next section is *needed* to generate impossibility results for *non-trivial* population prospects, since the traditional adequacy conditions for risk-free population axiology say nothing whatsoever about how to rank non-trivial prospects.

---

[24]However, note that even in the non-risky case, Average Utilitarianism implies repugnant-like conclusions. See e.g., Arrhenius (2000a, forthcoming), Budolfson and Spears (2018).

4   Population State Dominance

As an illustration of a straightforward comparison of population prospects, consider prospects P_1 and P_2 in Table 1. We assume, in the below example (and indeed throughout this paper), that the only axiologically relevant features at stake are, first, the welfare of the people in the populations that the prospects might result in, and, second, the probabilities that the prospects result in these populations. Moreover, in the example below, we assume that the only person whose welfare P_1 and P_2 affect is Ann, who will lead a horrible and suffering life (denoted by "Ann: -100") if we choose P_2. The same is true if we choose P_1 and state of the world s_1 obtains. However, in state s_2 the choice of P_1 means that Ann will lead an excellent life (which we denote by "Ann: 100").

	s_1	s_2
P_1	Ann: -100	Ann: 100
P_2	Ann: -100	Ann: -100

Table 1: Illustration of dominance relationship

It seems evident that P_1 is better than P_2. After all, Ann is sure to lead a suffering life given the choice of P_2, but she could be spared these sufferings and lead an excellent life given the choice of P_1.

The above reasoning is an instance of the well-known *State-Wise Dominance* principle from decision theory. The role of this principle, which is rarely questioned as a normative constraint, is to ensure that the ranking of prospects coheres with the ranking of the outcomes (in our case, the populations) that the prospects could result in.[25] We see no reason why one would reject this as an adequacy condition for population prospect axiologies, assuming, as we have done, that the only axiologically relevant feature at stake is the welfare of the people involved and the probabilities of the possible populations (more on this below).[26] Nevertheless, one of the results we present in the next section does not assume State-Wise Dominance in full generality; and one result does not assume the principle at all.

---

[25]The classical theories of rational choice under uncertainty—such as Ramsey's (1926/1990), Savage's (1954), and Jeffrey's (1965)—all impose State-Wise Dominance for the type of situations we are considering, i.e., situations where the states' probabilities are unaffected by the agents' choice.

[26]It might be worth mentioning that unlike *Event*-Wise Dominance, State-Wise Dominance does not rule out intuitively reasonable risk attitudes that that are inconsistent with standard expected utility theory, such as the so-called *Allais Preference* (named after Allais 1953).

Let us call the fully general implication that State-Wise Dominance has for population prospects *Population State Dominance*, and formulate it thus:

Population State Dominance. *If for all $s_i \in S$, if $P_2(s_i) \precsim P_1(s_i)$, then $P_2 \precsim P_1$; and if, in addition, for at least one non-null[27] s_j, $P_2(s_j) \prec P_1(s_j)$, then $P_2 \prec P_1$.*

It might be worth noting that just as State-Wise Dominance, Population State Dominance entails that the value of an outcome (i.e., population) cannot depend on what other outcomes (i.e., populations) the prospect might result in, assuming that the outcome has been described in sufficient detail for the principle to be applicable. In other words, the principle rules out value dependencies across different states of the world.[28] This fact might give rise to an (ultimately misguided) objection to this constraint in the present context. Suppose populations **x** and **y** consist of different but equally many people, all having very high welfare; more generally, suppose we deem **x** and **y** to be equally good in all respects. And imagine that we are comparing a prospect that will result in **x** for sure (i.e., the prospect results in **x** in all states of the world) with a prospect that that has, say, a 0.5 probability of resulting in **x** and a 0.5 probability of resulting in **y** (i.e., the prospect results in **x** in some states of the world but results in **y** in other states). Then many people have the intuition that the second prospect is strictly better than the first, which might seem to violate Population State Dominance. After all, the second prospect but not the first gives the **y**-people a chance to exist and lead lives worth living.[29]

Should we give up Population State Dominance in response to the above objection? Perhaps this is a reason to think that the principle does not hold in full generality. But recall that we are assuming that the only axiologically relevant features at stake are, first, the welfare of the people in the populations that the prospects might result in, and, second, the probabilities that the prospects result in these populations. Now, if there is value to giving a person a chance of existing with positive welfare that is not exhausted by what is good for the person in question, then it is *not* true that the only axiologically relevant features at stake are the welfare of the people in the possible populations and the probabilities of these populations.

In any case, in the next section, we shall first assume that Population State Dominance

---

[27]A state is non-null if it has some probability of obtaining (in this case, according to the social planner).

[28]Note however that the fact that knowledge of what could have been can affect people's welfare is not an argument against Population State Dominance, since such effects would be reflected in the individual welfare measure.

[29]Diamond (1967) uses an example with this structure to criticise the much stronger Sure Thing Principle. For further discussion, see e.g. Broome (1991) and Stefánsson (2015).

holds whenever the only axiologically relevant features at stake are the two features discussed above. But we then present a result that limits the principle even further. And we end by presenting a result that does not assume dominance at all.

5   Population Prospect Impossibility Theorems

To answer the general question about the possibility of an adequate population prospect axiology, we can exploit the generally accepted and much studied adequacy conditions for (risk-free) population axiology, and add to them Population State Dominance, which, as previously mentioned, we take to be an adequacy condition for population prospect axiology. Although the traditional (risk-free) adequacy conditions are known to be mutually inconsistent as conditions on how to rank outcomes, they are not mutually inconsistent as conditions on how to rank *non-trivial* prospects, since none of them says anything about how to rank such prospects. Hence, a condition like Population State Dominance is needed to connect the ranking of outcomes to the ranking of non-trivial prospects.

We shall focus on three adequacy conditions for (risk-free) population axiology that generate a very simple and easy to demonstrate impossibility theorem, which we will extend to population prospect axiology. Some of these conditions can be and have been questioned. However, as will become evident, the various other impossibility theorems that have been proven with logically weaker and intuitively more compelling adequacy conditions for population axiology (see, e.g., Arrhenius, forthcoming) can be extended to population prospect axiology in a way analogous to how we extend the simple theorem. So, the aim here is not to prove an impossibility theorem involving only intuitively compelling adequacy conditions. The aim is instead to describe a recipe for translating *any* theorem for (risk-free) population axiology into a theorem for prospect population axiology on a domain that contains *only* non-trivial prospects.

Below we make use of one more piece of notation. We stipulate that there is a $\gamma \in \mathbb{R}^+$—which can be arbitrarily close to 0—such that for any $x, y \in \mathbb{R}^+$ if $y < x$ but $x - y \geq \gamma$, then y is "slightly lower" than x. The three adequacy conditions for risk-free population axiology that we focus on are:

Egalitarian Dominance. *For any $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$, if $n(\boldsymbol{x}) = n(\boldsymbol{y})$ and for all $i, j \in \mathbb{N}$, $y_i < x_j$, then $\boldsymbol{y} \prec \boldsymbol{x}$.*

Quantity. *For any $x, y \in \mathbb{R}^+$ such that y is slightly lower than x, and for any $n \in \mathbb{N}$, there is a $m \in \mathbb{N}$, such that $m > n$, and such that $(x)_n \precsim (y)_m$.*

16

Quality. *There is a $x \in \mathbb{H}$ and a $n \in \mathbb{N}$, such that for any $y \in \mathbb{L}$ and any $m \in \mathbb{N}$, $(y)_m \precsim (x)_n$.*

Egalitarian Dominance is as uncontroversial as a principle of population ethics gets. Quality is a condition that rules out theories that imply the Repugnant Conclusion. Quantity may have less immediate intuitive appeal than the first two, but it should appeal to those who find some truth in the saying "the more good, the better".

Now, it turns out that the above seemingly week conditions are not mutually compatible, as illustrated by the following simple and well-known impossibility theorem (for risk-free population axiology):

Simple Impossibility Theorem (SIT). *No social welfare relation $\precsim$ on **X** satisfies Egalitarian Dominance, Quality, and Quantity.*[30]

To see why the above theorem is true, consider the sequence in Figure 1. We assume that well-being is perfectly equally distributed in each population. Moreover, assume that $\mathbf{x}_1$ in the diagram above is a population with very high welfare and that **y** is a population with very low positive welfare (the width of a block represents the number of people in the population, the height represents their welfare; dashes indicate that the block in question should really be much wider than shown). According to Quantity, there is a population $\mathbf{x}_2$ with slightly lower welfare than $\mathbf{x}_1$ which is at least as good as $\mathbf{x}_1$; a population $\mathbf{x}_3$ with slightly lower welfare than $\mathbf{x}_2$ which at least as good as $\mathbf{x}_2$; and so forth. Hence, we will finally reach population $\mathbf{x}_r$ with very low positive welfare.[31] Now, recall that we assume the social welfare relation to be transitive. Hence, $\mathbf{x}_r$ is at least as good as $\mathbf{x}_1$. But **y** is better than $\mathbf{x}_r$ according to Egalitarian Dominance. Hence, by transitivity, **y** is better than $\mathbf{x}_1$. However, by Quality, $\mathbf{x}_1$ is at least as good as **y** (for some $\mathbf{x}_1$ and **y** that can be used in an argument with the above structure). So, we get a contradiction: **y** is better than $\mathbf{x}_1$ which is at least as good as **y**. Hence, the Simple Impossibility Theorem: there is no social welfare relation—in other words, no population axiology—which satisfies Egalitarian Dominance, Quality, and Quantity.

---

[30]This theorem was first presented in Arrhenius (2000a). See also Arrhenius (2016)

[31]Here we exploit the fact that the formal framework we have been employing entails what has been called *Finite Fine-grainedness*: There exists a finite sequence of slight welfare differences between any two welfare levels. Finite Fine-grainedness ensures that one can get from one welfare level to another in a finite number of steps of intuitively slight welfare difference. Examples of such welfare differences could be some minor pain or pleasure or a shortening of life by a minute or two. For further discussion of Finite Fine-grainedness, and of possible theories of welfare that violate this assumption, see Arrhenius (2000a, 2005, 2016, forthcoming), Arrhenius and Rabinowicz (2015). For an interesting effort to challenge Finite Fine-grainedness (in light of the impossibility theorems in population ethics), see Carlson (nd) and Thomas (2018).
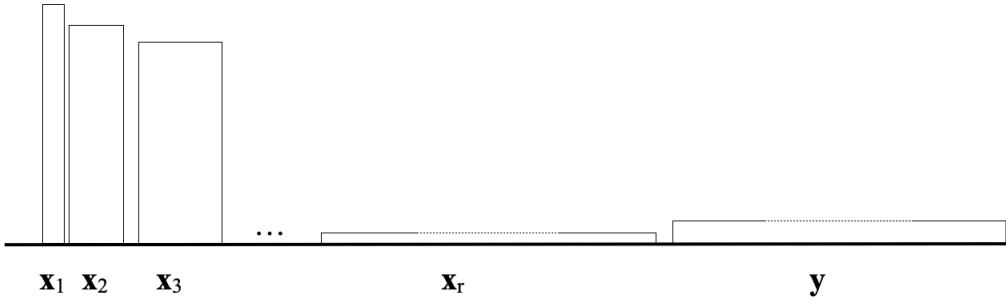
Figure 1: Illustration of SIT

The Simple Impossibility Theorem can easily be extended to the following impossibility result for population prospect axiologies, even if we limit the domain of the social welfare relation to prospects that are non-trivial (in the sense described above):

Simple Prospect Impossibility Theorem (SPIT)**.** *No social welfare relation $\precsim$ on $\boldsymbol{P}_{nt}$[32] satisfies Egalitarian Dominance, Quality, Quantity, and Population State Dominance.*

	s_1	s_2
P_1	$\mathbf{z}$	$\mathbf{x}_1$
P_2	$\mathbf{z}$	$\mathbf{x}_2$
...	...	...
P_r	$\mathbf{z}$	$\mathbf{x}_r$
P_{r+1}	$\mathbf{z}$	$\mathbf{y}$

Table 2: Prospects that prove SPIT

The prospects in Table 2 establish SPIT. P_1 to P_{r+1} are different population prospects, whereas s_1 and s_2 are different states of the world that determine which population results from the choice of each prospect. We assume that P_1 to P_{r+1} are non-trivial, meaning that the probabilities of states s_1 and s_2 are strictly between zero and one. (Those who find it hard to imagine that there could be *only* two distinct non-null states can interpret s_1 and s_2 as two events, each of which happens to resolve any uncertainty about which population results from prospects P_1 to P_{r+1}.) Now, assume that $\mathbf{x}_1$, $\mathbf{x}_2$, etc., and $\mathbf{y}$ in Table 2 are the populations depicted in Figure 1. So, for each $\mathbf{x}_i$, $\mathbf{x}_{i+1}$ is at least as good as $\mathbf{x}_i$ according

---

[32]Of course, there is no such relation on the set of trivial prospects ($\mathbf{P} \cap \neg\mathbf{P}_{nt}$) either; but that we already knew from SIT.

18

to Quantity, $\mathbf{y}$ is better than $\mathbf{x}_r$ according to Egalitarian Dominance, and thus $\mathbf{y}$ is better than $\mathbf{x}_1$ by transitivity. However, $\mathbf{x}_1$ is at least as good as $\mathbf{y}$ according to Quality. But that means that according to Population State Dominance, P_{r+1} is better than P_1 and P_1 is at least as good as P_{r+1}, irrespective of what population $\mathbf{z}$ represents. So, again we reach a contradiction. Hence, the Simple Prospect Impossibility Theorem: there is no social welfare relation—and thus no population *prospect* axiology—on a domain of non-trivial prospects which satisfies Egalitarian Dominance, Quality, Quantity, and Population State Dominance. (It is straightforward to see how the result can be extended to multiple non-null states.)

Note that since both Quality and Quantity are stated with the weak preference relation—that is, they are claims about the "at least as good as" relation, to use axiological terms—we need a version of Population State Dominance that applies dominance reasoning even when there is only "weak" state-wise dominance (i.e., where there is only weak preference for one prospect over another given some state). However, by strengthening Quality, we can make do with a weaker dominance condition that only applies dominance reasoning when there is strict state-wise dominance.[33]

We could of course also replace Population State Dominance with *Stochastic Dominance*, which, in a state-space framework like ours, entails Population State Dominance. Finally, it might also be worth pointing out that the assumption that the different prospects result in the *same* population if one of the two states obtains is neither essential to the above result nor the next theorem we introduce.

Before introducing the next theorem, it might be worth reminding the reader of a point that we previously made, namely, that prospect impossibility results such as those proven above cannot be achieved by only using risk-free adequacy conditions, such as Egalitarian Dominance, Quantity, and Quality. For while these risk-free conditions tell us how to rank *populations* (and hence how to rank *trivial* population prospects), they tell us nothing about how to rank *non-trivial* prospects that could result in different populations. For instance, while Egalitarian Dominance, Quality, and Quantity together entail that $\mathbf{y}$ is better than $\mathbf{x}_1$ and that $\mathbf{x}_1$ is at least as good as $\mathbf{y}$, that does not suffice to get an impossibility result for non-trivial prospects, since this non-transitive outcome-ranking by itself does not preclude that there exists a consistent (in particular, transitive) ranking of *non-trivial prospect*.

More generally, to derive an impossibility result for non-trivial prospects, one either has to use a condition like Population State Dominance to connect the ranking of outcomes to the ranking of prospects, or one has to reformulate the standard conditions as prospect

---

[33]We thank Dean Spears for encouraging us to clarify this.

conditions (a possibility we explore blow). For instance, while Egalitarian Dominance, Quantity, and Quality are mutually inconsistent as conditions on how to rank outcomes, these three conditions are by themselves mutually *consistent* as conditions on how to rank non-trivial prospects. After all, conditions that say *nothing* about how to rank non-trivial prospects are obviously mutually consistent as conditions on how to rank such prospects.

Let's now consider the second approach mentioned above, by reformulating Egalitarian Dominance, Quality, and Quantity as prospect conditions by building Population State Dominance into the original conditions. Consider now the following conditions:

Egalitarian Prospect Dominance. *For any $x, y, z \in X$ such that $n(x) = n(y)$ and $y_i < x_j$ for all $i, j \in \mathbb{N}$, for any $E \subset E$, and for any $P, Q \in P$, if $P(\neg E) = x$ and $P(E) = z$ while $Q(\neg E) = y$ and $Q(E) = z$, then $Q \prec P$.*

Quantity Prospect Dominance. *For any $z \in X$ and for any $x, y \in \mathbb{R}^+$ such that y is slightly lower than x, and for any $n \in \mathbb{N}$, there is a $m \in \mathbb{N}$, such that $m > n$, and such that for any $E \subseteq E$ and for any $P, Q \in P$, if $P(E) = (x)_n$ and $P(\neg E) = z$, while $Q(E) = (y)_m$ and $Q(\neg E) = z$, then $P \precsim Q$.*

Quality Prospect Dominance. *There is a $x \in \mathbb{H}$ and a $n \in \mathbb{N}$, such that for any $y \in \mathbb{L}$ and any $m \in \mathbb{N}$, for any $z \in X$, for any $E \subseteq E$, and for any $P, Q \in P$, if $P(E) = (x)_n$ and $P(\neg E) = z$, while $Q(E) = (y)_m$ and $Q(\neg E) = z$, then $Q \precsim P$.*

Now, it should be straightforward to see that even if the domain is restricted to non-trivial prospects, that is, prospects that confer a probability strictly between zero and one on different populations, the above three conditions form an inconsistent set. That is, we can prove:

Pure Prospect Impossibility Theorem (PIT). *No social welfare relation $\precsim$ on P_{nt} satisfies Egalitarian Prospect Dominance, Quality Prospect Dominance, and Quantity Prospect Dominance.*

The prospects in Table 2, for instance, establish the Pure Prospect Impossibility Theorem. For notice that by transitivity, Egalitarian Prospect Dominance, Quality Prospect Dominance, and Quantity Prospect Dominance result in an inconsistent ranking of the prospects in Table 2, in exactly the same way as the original Egalitarian Dominance, Quality, Quantity and Population State Dominance resulted in an inconsistent ordering of the prospects in Table 2.

One advantage of PIT, as compared to SPIT, is that PIT does not assume Population State Dominance to always hold. So, in particular, those who want to resist Population State

Dominance due to complementarities between outcomes in different states of the world, as discussed above, should be more troubled by PIT, which only assume that Population State Dominance holds with respect to the populations involved in the conditions.

However, in case someone objects to *any* appeal to Population State Dominance, even a restricted one as used in the Pure Prospect Impossibility Theorem, it might be worth noting that impossibility results can be generated without any such appeal, albeit with more complex conditions.[34]

Consider now the following two conditions:

Probability-Quantity Tradeoff. *For any $m, n \in \mathbb{N}$, for any $x \in \mathbb{H}$ and any $y \in \mathbb{L}$, for any $E \in \mathbf{E}$ such that $p(E) = q > 0$ and for any $P \in \mathbf{P}$ such that $P(E) = (x)_m$ and $P(\neg E) = (y)_n$, there is some $F \in \mathbf{E}$, which possibly partly overlaps with E, such that $p(F) = q' < q$, and some $o \in \mathbb{N}$, such that if $Q(F) = (x)_m$ and $Q(\neg F) = ((y)_n, (y)_o)$, then $P \precsim Q$.*

Probability-Quality Tradeoff. *There is a $x \in \mathbb{H}$ and a $n \in \mathbb{N}$, such that for any $k, l \in \mathbb{N}$, and for any $y \in \mathbb{L}$, there are some (possibly overlapping) events, $G, H \in \mathbf{E}$, such that $p(H) < p(G)$, such that if $P(G) = (x)_n$ and $P(\neg G) = (y)_k$, while $Q(H) = (x)_n$ and $P(\neg H) = ((y)_k, (y)_l)$, then $Q \prec P$.*

The thought behind Probability-Quantity Tradeoff is that we can reduce the probability that a prospect results in the population $(x)_m$, where people have very high welfare, without making the resulting prospect worse, as long as we sufficiently increase the number of people in the alternative population that the prospect results in if it does not result in $(x)_m$. In contrast, the thought behind Probability-Quality Tradeoff is that there is a limit to how much average welfare we are willing to risk foregoing for the sake of a chance at higher total welfare. As can easily be verified, Prospect Average Utilitarianism violates Probability-Quantity Tradeoff whereas Prospect Total Utilitarianism violates Probability-Quality Tradeoff.

To see how Probability-Quantity Tradeoff and Probability-Quality Tradeoff jointly lead to inconsistency, given transitivity, suppose we start with a probability p that is close to 1. (To make the argument as simple as possible, we replace the events/states by their probabilities.) By Probability-Quantity Tradeoff, there is some quantity δ that we could subtract from p such that the following holds. Suppose that prospect P_0 has probability p of resulting in a population $(x)_m \in \mathbb{H}^m$ and probability $1 - p$ of resulting in some population $(y)_n \in \mathbb{L}^n$. Moreover, suppose that prospect P_1 has probability $p - \delta$ of resulting

---

[34]We thank Tim Campbell for suggesting we add a result like this to the paper.

in population $(x)_m$ and probability $1 - (p - \delta)$ of resulting in $((y)_n, (y)_o)$. Then P_1 is no worse than P_0. But continuing this reasoning, we can generate a sequence of prospects, by incrementally increasing the quantity that is subtracted from p, until we reach a quantity γ for which the following holds. Suppose that P_n has probability $p - \gamma$ of resulting in $(x)_m$ and probability $1 - (p - \gamma)$ of resulting in some population $((y)_n, (y)_o, ...(y)_p)$, which is much bigger than $(y)_n$. Then P_n is no worse than the immediately preceding prospect in the sequence. Hence, by transitivity, P_n is no worse than P_0. Since we get this result irrespective of the welfare (and size) of the $(x)_m$ population, and irrespective of p and γ, we have a violation of Probability-Quality Tradeoff.

6   Concluding Remarks

We have found that there can be no adequate axiology for non-trivial population prospects, since no such axiology can satisfy conditions that, we contend, any adequate population prospect axiology should satisfy. Although we have focused only on one of the many impossibility theorems that have been proven for risk-free population axiology, it should be evident that the other theorems can be extended to population prospect axiology in an analogous way to how we extended the Simple Impossibility Theorem, that is, by exploiting Population State Dominance either as a general condition or in a restricted form built into the adequacy conditions.

One might have hoped that the impossibility theorems from risk-free population ethics could be avoided by removing some of their idealising assumptions, for instance, the assumption that we know for sure which population an alternative will result in. But, as we now have seen, removing the idealising assumption of certainty will not help us avoid analogous impossibility results for non-trivial population prospects.

References

Allais, M. 1953. "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine." *Econometrica*, 21(4): 503–546.

Arrhenius, Gustaf. 2000a. *Future Generations: A Challenge for Moral Theory*: Uppsala University Printers.

——— 2000b. "An impossibility theorem for welfarist axiologies." *Economics & Philosophy*, 16(2): 247–266.

——— 2003a. "Feldman's Desert-Adjusted Utilitarianism and Population Ethics." *Utilitas*, 15(2): 225.

——— 2003b. "The very repugnant conclusion." *Uppsala Philosophical Studies*, 51.

——— 2005. "Superiority in Value." *Philosophical Studies*, 123(1-2): 97–114.

——— 2016. "Population Ethics and Different-Number-Based Imprecision." *Theoria*, 82(2): 166–181.

Arrhenius, Gustaf and Wlodek Rabinowicz. 2015. "Value Superiority." in Iwao Hirose and Jonas Olson eds. *The Oxford Handbook of Value Theory*: Oxford University Press: 225–248.

Arrhenius, Gustaf. forthcoming. *Population Ethics: The Challenge of Future Generations*: Oxford University Press.

Arrhenius, Gustaf, Jesper Rydberg, and Torbjörn Tännsjö. 2017. "The Repugnant Conclusion." in Edward N. Zalta ed. *The Stanford Encyclopedia of Philosophy*: Metaphysics Research Lab, Stanford University, spring 2017 edition.

Asheim, Geir and Stéphane Zuber. 2014. "Escaping the repugnant conclusion: Rank-discounted utilitarianism with variable population." *Theoretical Economics*, 9(3): 629–650.

——— 2016. "Evaluating intergenerational risks." *Journal of Mathematical Economics*, 65(C): 104–117.

Bell, David. 1982. "Regret in Decision Making Under Uncertainty." *Operations Research*, 30: 961–81.

Blackorby, Charles and David Donaldson. 1984. "Social criteria for evaluating population change." *Journal of Public Economics*, 25(1-2): 13–33.

Blackorby, Charles, Walter Bossert, and David J Donaldson. 2005. *Population issues in social choice theory, welfare economics, and ethics*: Cambridge University Press.

Bossert, Walter. 2017. "Anonymous welfarism, critical-level principles, and the repugnant and sadistic conclusions." working paper, University of Montréal.

Broome, John. 1991. *Weighing Goods: Equality, Uncertainty and Time*: Wiley-Blackwell.

——— 1992. *Counting the Cost of Global Warming*: Strond: White Horse Press.

——— 2004. *Weighing Lives*: Oxford University Press.

——— 2010. "The Most Important Thing About Climate Change." in Jonathan Boston, Andrew Bradstock, and David Eng eds. *Public Policy: Why Ethics Matters*: ANU E Press: 101–16.

——— 2015. "Climate Change: Life and Death." in Jeremy Moss ed. *Climate Change and Justice*: Cambridge University Press: 184–200.

Buchak, Lara. 2013. *Risk and Rationality*: Oxford University Press.

Budolfson, Mark and Dean Spears. 2018. "Why the repugnant conclusion is inescapable." working paper, Princeton University CFI.

Carlson, Erik. nd. "On Some Impossibility Theorems in Population Ethics." working paper, Uppsala University.

Diamond, Peter. 1967. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: Comment." *Journal of Political Economy*, 75.

Diecidue, Enrico and Jeeva Somasundaram. 2017. "Regret theory: A new foundation." *Journal of Economic Theory*, 172(C): 88–119.

Feldman, Fred. 1997. *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy*: Cambridge University Press.

Glover, Jonathan. 1977. *Causing Death and Saving Lives*: Penguin Books.

Hurka, Thomas. 1983. "Value and population size." *Ethics*, 93(3): 496–507.

Jeffrey, Richard. 1965. *The Logic of Decision*: The University of Chicago Press.

Loomes, Graham and Robert Sugden. 1982. "Regret Theory: An Alternative Theory of Rational Choice Under Risk." *The Economic Journal*, 92: 805–824.

Nebel, Jacob M. 2017. "Priority, Not Equality, for Possible People." *Ethics*, 127(4): 896–911.

——— 2019. "An Intrapersonal Addition Paradox." *Ethics*, 129(2): 309–343.

Ng, Yew-Kwang. 1989. "What Should We Do About Future Generations?: Impossibility of Parfit's Theory X." *Economics & Philosophy*, 5(2): 235–253.

Parfit, Derek. 1984. *Reasons and Persons*: Oxford.

——— 2016. "Can we avoid the repugnant conclusion?" *Theoria*, 82(2): 110–127.

Pivato, Marcus. 2020. "Rank-additive population ethics." *Economic Theory*, 69(4): 861–918.

Quiggin, John. 1982. "A theory of anticipated utility." *Journal of Economic Behavior and Organization*, 3(5): 323–343.

Ramsey, Frank P. 1926/1990. "Truth and Probability." in D. H. Mellor ed. *Philosophical Papers*: Cambridge University Press.

Roberts, M. A. 2007. "The Non-Identity Fallacy: Harm, Probability and Another Look at Parfit's Depletion Example." *Utilitas*, 19(3): 267–311.

Savage, Leonard J. 1954. *The Foundations of Statistics*: Wiley Publications in Statistics.

Sider, Theodore R. 1991. "Might theory X be a theory of diminishing marginal value?" *Analysis*, 51(4): 265–271.

Spears, Dean and Mark Budolfson. 2019. "Why Variable-Population Social Orderings Cannot Escape the Repugnant Conclusion: Proofs and Implications." Discussion Paper 12668, IZA.

Stefánsson, H. Orri. 2015. "Fair Chance and Modal Consequentialism." *Economics and Philosophy*, 31(3): 371–395.

Stefánsson, H. Orri and Richard Bradley. 2015. "How Valuable Are Chances?" *Philosophy of Science*, 82(4): 602–625.

——— 2019. "What Is Risk Aversion?" *British Journal for the Philosophy of Science*, 70(1): 77–102.

Thomas, Teruji. 2018. "Some Possibilities in Population Axiology." *Mind*, 127(507): 807–832.

Tversky, Amos and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty*, 5(4): 297–323.

von Neumann, John and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*: Princeton University Press.

Voorhoeve, Alex and Marc Fleurbaey. 2016. "Priority or Equality for Possible People?" *Ethics*, 126(4): 929–954.

Zuber, Stéphane and Geir B Asheim. 2012. "Justifying social discounting: the rank-discounted utilitarian approach." *Journal of Economic Theory*, 147(4): 1572–1601.